Glossary

archive

A file on storage media is contained in an archive. Enstore restricts archives to one file each. Therefore, in enstore, the terms file and archive may essentially be used interchangeably. The archive consists of the file (the data) plus some information that is wrapper-dependent. In the case of cpio wrappers, for example, an archive consists of a cpio header, the file, a cpio trailer (to make the file divisible by 512), and a filemark.

crc

An Adler 32 bit crc polynomial is used to verify that data has been stored properly (it's used like a checksum, but is less prone to multiple-bit errors). When a transfer is complete, both sides calculate the crc and compare the values.

cwd

current working directory

dCache

DiskCache, or dCache, is a data file caching system which acts as an intelligent manager between the user and the data storage facilities. It optimizes the location of staged copies according to an access profile. DCache is designed for use by off-site users via ftp. It decouples the (potentially slow) network transfer rate from the (fast) storage media I/O rate in order to keep the enstore system from bogging down.

door (for dCache)

A door is an instance of the dCache software installed on a server. Each door is associated with a particular port on the server, and has its own access profile.

encp

Encp is an end-user command, designed to be used with enstore, that is used to copy data files from disk to storage media and vice-versa. Encp communicates file family and other data file configuration information between enstore and pnfs.

enstore

Enstore is the mass storage system implemented at Fermilab as the primary data store for experiments' large data sets. It provides distributed access to data on tape or other storage media both locally and over networks.

file family

A file family is an arbitrary ASCII name that defines a category, or family, of data files. Each experiment defines a set of file families for its data. A given storage volume may only contain files belonging to one file family.

file family width

File family width is an integer value used to limit write-accessiblity on data storage volumes. At any given time, enstore limits the number of volumes associated with a given file family that are open for writing to the value of the file family width.

file family wrapper

A file family wrapper consists of information that gets added to the front and back of data files as they're written to media, and defines the files' format on the storage volume.

filemark

A filemark is a physical mark on tape indicating end of file. Tape drives recognize it and can do high speed searches over it. It is a physically different "writing" on the tape than a data EOF.

Kerberized ftp client

A Kerberized ftp client is an ftp client that implements Kerberos v5 authentication.

library

A library in Enstore is comprised of both the physical data storage media and a robot arm used to mount the media in attached drives. An Enstore library is typically called a robot.

library manager

A Library Manager (LM) is a software module (a server) which controls a virtual library. LMs receive requests for file copy jobs from users via **encp** and they distribute the jobs to the Movers.

pnfs namespace

Pnfs is an independent namespace package, originally written at DESY. It presents a collection of library database entries as a UNIX-like file system, and thus allows users to browse stored files as though they reside in this file system. Pnfs is mounted like NFS, but it is a virtual file system only. It maintains file grouping and structure information via a set of tag files in each directory. The **encp** product communicates this information between PNFS and the enstore servers when it uploads or downloads a data file.

pnfs tag files

See tags.

storage group

A storage group is an area in `/pnfs` namespace assigned by the enstore administrators to a particular experiment. The storage group name for an experiment is the experiment's top level directory under `/pnfs`.

storage volume

A unit of mass storage, e.g., a tape or a CD.

streaming (of files on tape)

Streaming refers to the sequential access of adjacent files on tape.

striping (of files on tape)

Striping refers to single files (usually large ones) being split onto two or more volumes, each writing simultaneously, in order to expedite the writing process. (Striping is not supported under enstore.)

virtual library

A Virtual Library (VL) is a subset of an Enstore data storage library. It can contain one and only one type of media. It is paired with its own library manager which controls it.

volume

See storage volume.

volume family

The triplet “storage group + file family + file family wrapper” is called a volume family. In order for different data files to be stored on the same volume, all three of these pnfs tags for the files must match.

wrapper

See file family wrapper.